the global scanner
a quarterly newsletter from the GES DISC

Summer 2001 ⁂ Volume 3 ⁂ Number 1

Processing the MODIS–Terra Consistent Data Products Year

a report by Lee Kyle, Bruce Vollmer, and Gary Roth

June saw the beginning of the processing of a consistent data year (11/2000–10/2001) of MODIS-Terra science products, which is scheduled to be completed by the end of November 2001. For this effort the MODIS Science Team delivered revised (MODIS Version 3) science algorithms that will improve the accuracy of the science products. At the GES DAAC a two-stream processing effort is underway. The direct processing of the incoming satellite data will continue on the ECS (EOSDIS Core System) processing system for the data taken on and after day 144 (May 24), 2001. A parallel S4PM (GES DAAC Simple, Scalable, Script-Based Science Processor-MODIS) processing system will handle the reprocessing for data days 305 (October 30), 2000, through 143, (May 23), 2001. This parallel processing effort should more than double the rate at which MODIS data products are produced. The MODIS Science Team is prioritizing the data to be reprocessed to Version 3, so that the reprocessing is not being done in strict chronological order. The Goddard Space Flight Center Earth Science (GES) Distributive Active Archive Center (DAAC) made extensive preparations for this increased effort. It set up a new processing system to handle the reprocessing, revised and expanded its tape archive and data distribution system, integrated into both processing systems the Version 3 science algorithms, and tested the systems.

The GES DAAC produces the MODIS Level 1A and 1B products, which include Earth-located at-satellite calibrated radiances, plus a cloud mask and atmospheric temperature and moisture profiles. It archives these plus the Level 0 MODIS measurements, which are in engineering units and format. It also archives ancillary data and the higher Levels 3 & 4 ocean and atmospheric science products produced by the MODIS (science) Processing System (MODAPS). MODAPS, like the GES DAAC, set up a parallel processing system to handle reprocessing and took other steps in preparation for producing the consistent year Version 3 science products.

Satellite data products of the quality needed in many climate and global change studies generally go through several stages. For MODIS, these are

- Version 1, early release (beta) products, minimally validated and possibly containing significant errors
- Version 3, provisional products, algorithm validation is not complete; hence, product quality may not be optimal (please contact the MODIS Science Team before using these data in a publication)
- Validated MODIS products suitable for scientific uses and publications, although some future improvements may be made.

There were no MODIS Version 2 products. Some of the Version 3 products may be validated by the science team while other products may require improved algorithms and another reprocessing. The Version 3 provisional products will be produced with fixed algorithms for the period November 2000 through October 2001. The present plan is to complete the Version 3 consistent year processing by the end of November 2001.

The MODIS Science Team is responsible for constructing the science algorithms and judging the quality of the produced products; therefore, any questions concerning the algorithms used and data quality should be addressed to them. See

http://modis.gsfc.nasa.gov/MODIS/

The team used the Version 1 products to study the idiosyncrasies of the MODIS instrument on Terra and improve their original science algorithms. This spring they delivered improved science algorithms to the GES DAAC and MODAPS production facilities. At the GES DAAC the Science Software Integration and Test team led by Bruce Vollmer integrated the Level 1 and at-

continued on page 2

Processing the MODIS-Terra Consistent Year

continued from page 1

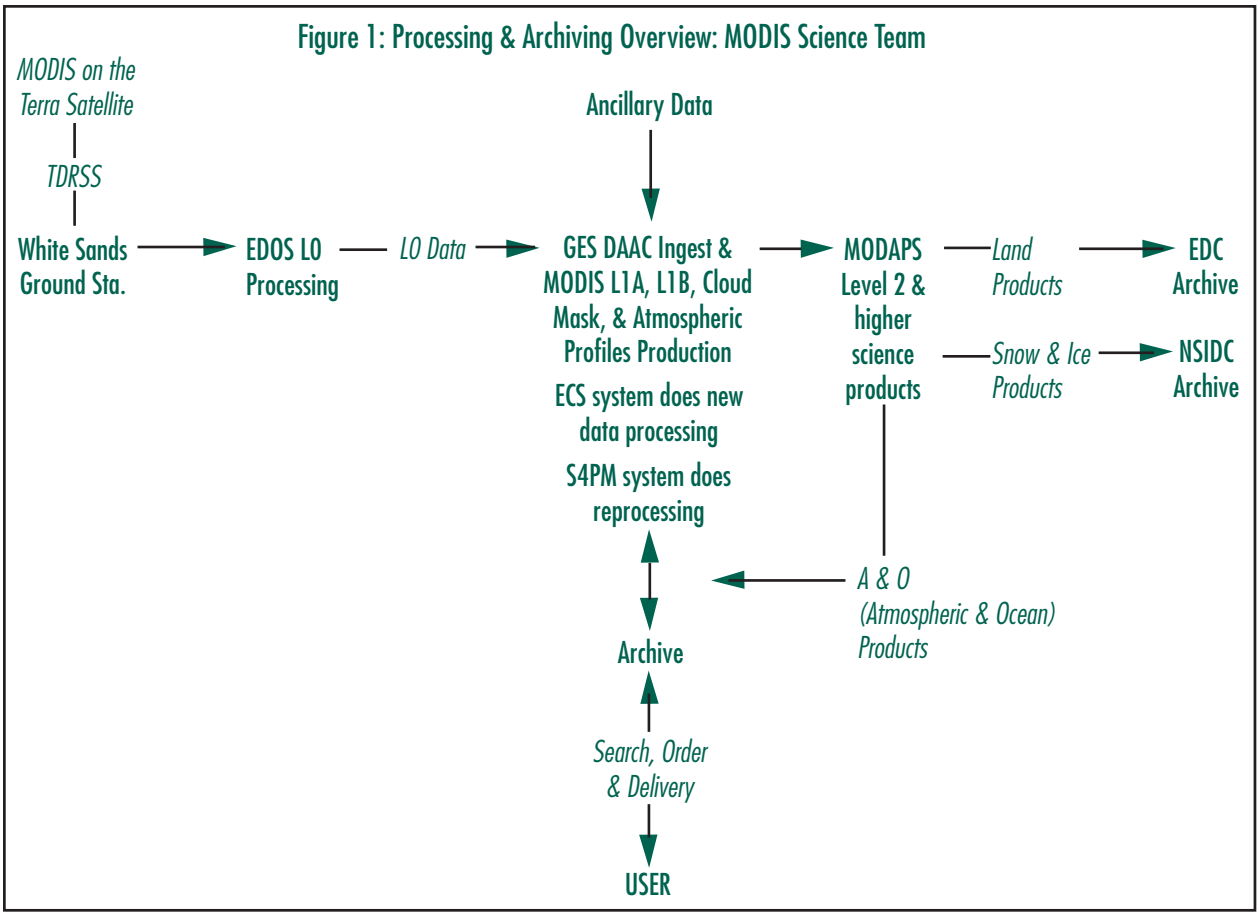
mospheric products science software packages into both the present ECS processing system, which processes the MODIS data stream coming from the satellite, and the new S4PM processing system. According to standard procedure, both systems were tested to ensure that the science algorithms were operating as expected by the science team. A test was also made to ensure that MODAPS and the GES DAAC could handle the sharply increased volumes of data that would pass back and forth during reprocessing.

Figure 1 is an overview of the MODIS-Terra data processing and archive system. The MODIS instrument on the Terra satellite monitors Earth's ocean, land, and atmosphere with 36 spectral bands ranging from ultraviolet to thermal infrared (0.405–14.385 μm). The Terra satellite is in a Sun synchronous orbit, which allows the MODIS sensors to cover the entire planet every 2 days. The 12-bit measurements are stored on the spacecraft's solid state recorder and later relayed to the White Sands ground station by the NASA Tracking and Data Relay Satellite System (TDRSS). White Sands transmits the data to the Earth Science Data Operations System (EDOS) that formats it into the MODIS L0 product and passes it along to the GES DAAC for archiving and further processing.

The GES DAAC does the preliminary science processing (Levels 1A & 1B plus some atmospheric products) and collects ancillary data needed to produce higher level science products. The original ECS processing system is used for this work. All of these data are archived, and all but the L0 data are passed on to MODAPS where the higher level science products (ocean productivity, land cover, atmospheric aerosols, etc.) are produced. According to type, these products are archived in various places. Ocean and atmospheric products come to the GES DAAC, land products go to EDC (Earth Resources Observation Systems [EROS] Data Center), and snow and ice products to NSIDC (National Snow and Ice Data Center). The MODIS Science Team has overall responsibility for the program and checks on everything from MODIS instrument behavior to the quality of the final products.

For reprocessing, the L0 and ancillary data are pulled out of the archive and sent through the S4PM system to produce the revised lower level science products that are then archived. They, together with the required ancillary data, are also sent to MODAPS for the

Figure 1: Processing & Archiving Overview: MODIS Science Team



production of revised higher level products. The S4PM processing system is easier to maintain and modify than ECS, and it is planned to eventually use S4PM systems to do all MODIS processing at the GES DAAC. For a brief description of the S4PM system see the article by Chris Lynnes in the *Global Scanner,* Vol. 2, No. 3-4. ECS will still be used to process future satellite data (e.g., from the Atmospheric Infrared Sounder [AIRS] on the Aqua satellite).

APPROXIMATE NORMAL (1X) PROCESSING DATA FLOW RATES

PRODUCT TYPE	SENT TO	RATE (GB/DAY)
GES DAAC Processing		
L0, incoming MODIS data in engineering units	Archive & Level 1 processing	84
L1 A & B plus other products	Archive and MODAPS	350
MODAPS Level 2 and Above Products		
Atmospheric and Ocean (chiefly ocean)	GES DAAC Archive	180
Land	EDC Archive	292
Snow and Ice	NSIDC Archive	15

Data Flow Rates

The table to the right indicates the normal processing rates for data flowing down from the satellite when 100% of the data from the MODIS instrument is being processed. These average rates are locally termed 1X. Since problems arise from time to time, the actual rates can be less (during problem periods) or greater (during catch-up periods). With reprocessing underway the average data flows are expected to be about 2.4X (1X in the ECS system and 1.4X in the S4PM reprocessing system). Reprocessing started June 22 with the data from March 5 (day 64), 2001, and will proceed through May 23 (day 143). Then it will jump back to October 30 (day 305), 2000, and work up through March 4 (day 63), 2001. The ECS system started processing using the consistent year science algorithms with data from May 24, 2001, MODIS data. A second version of the S4PM system took over forward processing of MODIS data from the ECS system in September, which will then be used for other, non-MODIS processing. The consistent year data processing is planned to be completed during November 2001. Reprocessing started at a 1X rate but will ramp up to 1.4X in order to complete the consistent year processing on schedule.

The first MODIS processing and archiving priority is Level 0 engineering measurements, since everything downstream depends on them. Next comes Levels 1A & 1B products, which include the vital Earth located, at-satellite calibrated radiances. Science products, the reason for the whole project, come last. But since their accuracy is critical, the accuracy of both the input data and science algorithms must be carefully checked. Thus production at MODAPS, the science processor, started at a slower pace than it did at the GES DAAC. During fall 2000 the GES DAAC archive received about 29 GB/day of MODIS ocean and atmosphere science products from MODAPS. This jumped to 109 GB/day for the data period (2/22/01–5/29/01). During this initial period many science products were not produced on a global scale. As indicated in the table, the MODAPS normal 1X processing rate is presently 487 GB/day, with 180 GB/day coming to the GES DAAC. With reprocessing running at 1.4X, their total rate will average 2.4X (1,169 GB/day). During catch-up periods processing will run considerably higher than this. The processing capabilities of MODAPS were recently expanded considerably to include separate direct and reprocessing systems. Edward Masuoka reports that the reprocessing system can maintain reprocessing rates up to 2.4X.

The GES DAAC V2 (ECS) Archive

The Terra MODIS data archived at the GES DAAC is stored in a tape archive that consists of large, computer controlled robotic tape silos. Incoming data files are sorted by their Earth science data types (ESDTs) into one of several tape volumes. There is a separate ESDT for each type of input and output file needed by the production processing generation executables (PGEs) that produce the MODIS science products. The files are not strictly time ordered but are stored on tape in the order that they are received and ingested into the system. When a tape is full, a new tape is pulled from a pool of preformatted tapes. The number of tapes in a volume grows as more product files come in. The ESDT to volume mapping is dependent on the volume of data (number of files times size of the files). The volumes are designed so that the data flow is balanced among the active silos to optimize their use. The computer system keeps track of the volume and tape location of each file so that they can be expeditiously retrieved. It also controls the robotic system that does the actual tape handling inside each silo.

The ECS archive started with two Storage Tek (STK) robotic silos, each of which holds 5000 tapes. D3 tapes were originally used but, as STK is phasing these out, the GES DAAC has switched to T9940 tapes, which the silos can also handle. Each T9940 tape has a capacity of 60-100 GB of data, depending on the data compression factor. The archive uses a hardware

continued on page 4

Processing the MODIS-Terra Consistent Year continued from page 3

compression method that automatically compresses data as they are written to tape and decompresses them as they are read off. Depending on the data type the compression ratio (native to compressed) ranges between 1.2 and 1.75, with the mean ratio estimated to be greater than 1.5. T9940 tapes hold slightly more data than the old D3 tapes and have similar data transfer, tape mount, and search speeds. The main benefit of 9940 drives is that they have a smaller footprint; 20 9940 drives fit in the same silo wall space as 4 D3 drives. The eight original D3 tape drives in Silo 1 were in two cabinets attached to the silo's walls. They were replaced by eight 9940 drives in one cabinet. The new 9940 drives can't read the old D3 tapes so the data on the old tapes are being migrated to the new T9940 tapes. The new Version 3 MODIS data are being written to T9940 tapes.

The present plan calls for six silos.
Silo I will contain 5000 T9940 tapes after migration is complete.
Silo II will contain 5000 T9940 tapes after migration is complete.
Silo III can hold 5000 tapes but currently has around 100 of the lower capacity 9840 tapes installed. (Browse data have only recently begun to be archived. We will add more tapes to this silo as demand increases. Each 9840 tape will hold up to 40 GB of data.)
Silo IV has the silo and internal robotic hardware in place but not yet integrated (i.e., no tapes, drives, SGI, nor Sun workstation interfaces).
Silos V and VI are to be developed.

Silo III contains "samples" of the archived data (i.e., browse products) that users of the GES DAAC Web interface can call up to see if they want to order either the full scale data products or the browse product itself. Browse products are reduced resolution images. The system attempts to fill these real time requests as quickly as possible. The 9840 medium was selected for this silo for faster access time. Since the average file size is small compared to the full resolution archive files, tape capacity was not an issue in selecting the medium.

Distributing ECS Data Products to Customers

The GES DAAC's capacity to distribute Earth science data is also being upgraded. Figure 2 shows an outline of the DAAC's customer data distribution system with new features indicated. Data ingest is the first archiving step. It registers each file with the GES DAAC with the file's description, its source, time, geographic location, etc., and determines where it will be stored. Data ingest is also used in the search and order programs customers use to order data. Some customers leave standing orders (subscriptions) for specific types of data over specified geographic areas. While all MODIS data are archived to tape, a parallel stream of subscription data is sent off as they come in. Thus subscription orders do not have to go through the additional step of retrieving the ordered data from tape. Data orders can be shipped either electronically or on solid media.

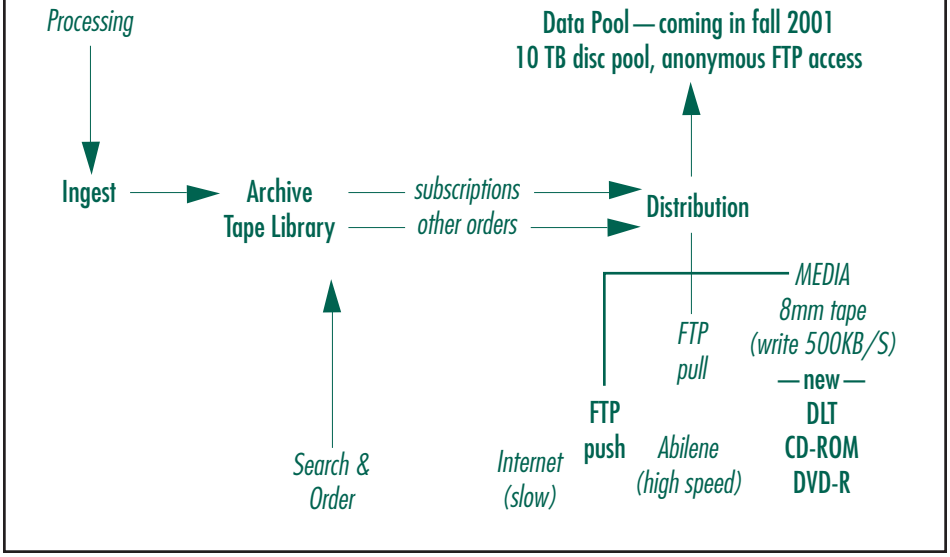
Solid media data shipment — Previously individual media orders went out only on tape, although some preselected data collections were available on CDs. Now individual media orders can also be shipped on DLT tapes, DVD-R and CD-R. The DLT tapes are useful for very large orders (a large fraction of a TB or more) while CDs are useful for orders of a few GB or less. 8-mm tapes and DVD-Rs are best for intermediate sized orders. Of course, the customer's facilities are an important factor in what medium is specified in the order.

Electronic media shipment — Presently data go out either by FTP push or FTP pull. FTP push is available over the Internet or over the Abilene high speed system for those with access to it. Abilene, an Internet2(R) backbone network, presently connects over 180 Internet2 universities and research laboratories in the United States and connects to other high speed networks in this country and internationally. This fall a new 10 TB disc data pool will come on line at the GES DAAC. It will be accessible by anonymous FTP and will have a connection to Abilene. In the coming year the capacity of the data pool will be increased and near-archive data mining and other services will be available on it.

For the foreseeable future the volume the GES DAAC can distribute to customers will have some limits. The present EOS project design, constrained by budget considerations, calls for a total distribution rate of 1X to

continued on page 7

Figure 2. The GES DAAC's Earth Science Customer Data Distribution System With New and Coming Improvements Indicated



The GES DAAC User Working Group

by Lee Kyle & Steve Kempler

The purpose of the Goddard Earth Science (GES) Distributed Active Archive Center (DAAC) User Working Group (UWG) is to provide guidance and recommendations that will better prepare the GES DAAC to serve the Earth science community in the future within the constraints of available DAAC resources. It consists of knowledgeable DAAC users in various fields of remote-sensing related science research and applications who are willing to review the GES DAAC's mode of operation and advise it on ways to best serve its customers. DAAC customers include the science teams responsible for providing data and the multitude of users who take data from the archive for research, applications, and education. The new UWG has held two meetings (November 14 & 15, 2000, and June 14 & 15, 2001) that have proved useful to GES DAAC management. This article briefly gives the background of the UWG and summarizes the subjects discussed at the second meeting.

Early in NASA's history Congress set up by law the National Space Science and Data Center (NSSDC) to archive NASA derived data sets. Over the past 3 decades the problem of global environment and climate change have become increasingly important, and the relevant data taken by NASA has increased exponentially. In 1992 the growing Earth science data collection was split off and established in a new, separate organization — the Earth Science Distributive Active Archive Centers (DAACs) — that consists of several data centers (currently eight) scattered around the United States, each specializing in different disciplines of Earth science data. NSSDC retains the planetary and astrophysical data collections. The GES DAAC specializes in atmospheric dynamics, atmospheric chemistry, hydrology, ocean color, and land biosphere data sets.

When the Earth science DAACs were established, NASA Headquarters formed short-term User Working Groups (UWGs) of interested climate researchers to help guide the efforts of the DAAC managers toward best serving the science community. The UWG at the GES DAAC faded away several years ago, while the number and types of customers increased dramatically. Issues regarding data management technology, NASA direction, and systems issues begged for a reinstatement of the GES DAAC UWG.

At the June UWG meeting DAAC managers reviewed the DAAC structure and operations (see "The GES DAAC, How It Works," in the *Global Scanner,* Vol. 1, No. 1) and pointed out recent improvements and continuing problems. UWG members discussed these points and also introduced questions and concerns of their own. Dr. Stan Morain, of the University of New Mexico, serving as interim chair of the UWG, led the discussions.

UWG members generally agreed that the GES DAAC was doing a good job with the V0 and V1 archive systems. These handle data from Tropical Rainfall Measuring Mission (TRMM), the Sea-Viewing Wide Field-of-View Sensor (SeaWiFS), and other missions previous to the launch of the Terra satellite. Considering the resources available, they also thought that the GES DAAC was doing reasonably well with the V2 archive system, which processes and archives data from the Moderate-Resolution Imaging Spectroradiometer (MODIS) on the Terra satellite. The V2 system is also known as the ESDIS Core System (ECS). Some of the more important items discussed were

- ***Processing and archiving speeds*** — Last year the V2 system had the ability to process and archive 100% of MODIS and associated data sent to it. This is termed a 1X capability. It included the ability to hit higher short-term processing rates to make up for periods of delayed data arrival or limited reprocessing to correct for processing errors. To handle reprocessing, the V2 system was upgraded to 240% (2.4X) of its original speed, seeing peaks of 300% (3X). Dr. Vincent Salomonson, the MODIS-T Team Leader, said that a higher sustainable processing and archiving rate was desirable for both present and future reprocessing programs.
- ***Limited ability to fill customer orders*** — The V2 system, delivered by Raytheon on contract to the Earth Sciences Data Information System (ESDIS) project, was designed to be able to send data out to its customers at the 1X rate. Because of resource constraints the project put highest priority on getting quality data into the archive. It is presumed that services to customers will later be improved. Steve Kempler reported the DAAC has received many small orders as well as some individual orders for all of the MODIS data. The latter were turned down since filling one such order would block out all other customers. So far there have not been enough small customers to strain the system, but this is expected to change. An informal survey indicated that potential users are aware of MODIS data, but are concerned about data maturity, had difficulties with the ordering system, were unable to download the large MODIS archive files, or were unfamiliar with the HDF-EOS format used in the DAACs. Data reprocessing will take care of the data maturity concern, while the DAAC is working to minimize the other problems. To meet the reported problems the GES DAAC will be pro-

continued on page 6

The GES DAAC User Working Group
continued from page 5

viding regional subsets for those not interested in complete MODIS archive files, and has in place reformatting tools for those who are unfamiliar with the HDF-EOS archive format. This is particularly important when new versions of HDF-EOS are introduced. The GES DAAC is also taking steps to allow customer designed data subsetting and data mining on the GES DAAC facilities. For additional information about these and other steps taken to improve customer access to MODIS data see the companion article in this issue, "Processing the MODIS-Terra Consistent Data Products Year."

- ***The difficulty in introducing new Earth science data types (ESDT) into the ECS processing system*** — This issue was raised by Wayne Esaias, leader of the MODIS ocean science team. At present it takes too long. Each type of input or output file used in the ECS processing system has a separate ESDT, metadata, and identification tag used to keep track of the files by the processing, archiving, and customer ordering systems. An ESDT may have between 50 and 200 attributes. While performing data quality control the science team needs to be able to quickly introduce new ESDTs or modify old ones. Introduction of new ESDTs is handled by the ESDIS project and not the GES DAAC. In the June 2001 meeting Mike Moore of the ESDIS project said that steps are being taken to make the introduction of new ESDTs easier and faster. The time to implement a new ESDT has now been cut from 8 or 9 months to 1 or 2 months. When asked, Moore indicated that to further speed up the process a reengineering of the ESDT implementation system is being considered.

Both data archiving and data processing procedures are rapidly evolving fields. Thus it behooves the GES DAAC management and the UWG to consider possible future developments. Among the items discussed were

- ***Long-term archiving*** — The UWG agreed that there is not at present an adequate program for the long-term archiving of climate data. This was discussed in both the November and June meetings. The NASA DAACs have responsibility for the climate data sent to them for the life of a mission plus 2 years. Mission life is normally defined as ending when data stop flowing. NOAA has the responsibility of being the long-term archive for climate data. However, in the past it has not been adequately funded to make all of these data easily accessible to the public. The increasing flood of good quality climate data will exacerbate this problem. The UWG agreed that it didn't matter to them where the data are stored as long as they are easily available. This problem should be kept before both the scientific community and the government until some reasonable program is set in place. It also advised that the GES DAAC tell the public whenever it removes a previously available data set and clearly explain what has happened to it and why.
- ***The upcoming NPOESS Preparatory Project (NPP) mission*** — In the future NASA plans that much of the long-term climate and global change data presently being measured by NASA Earth Science Enterprise (ESE) satellites such as Terra will be obtained from the planned new NOAA National Polar-Orbiting Operational Environmental Satellite System (NPOESS). The future NPOESS will be a merger of the NOAA and Department of Defense polar orbiting satellite programs. The present plan is to launch the NPOESS C1 satellite about 2008, and this NPOESS satellite series will continue to operate to about 2018. It will be principally an operational system with measurement products (environmental data records, EDR) designed to be available to meteorologists and other users within 20 to 90 minutes. The requirement for fast delivery will prevent many EDRs from being accurate enough for long-term climate studies. NASA will design a program and processing system to produce the more accurate climate data records (CDR) from the NPOESS measurements. Here accuracy and not speed has the highest priority. NASA and NOAA are cooperating on the NPP mission. It has two objectives: 1, continuation of the environmental measurement series started by the NASA Terra and Aqua satellites until the NPOESS C1 satellite is in operation; and 2, risk reduction for the NPOESS system by prototyping some of the instruments, the operational and science algorithms, and the processing systems required by NPOESS. The NPP satellite is planned for launch in 2005. The NPP program is being managed by the Goddard Space Flight Center. A system to process and archive the data will be developed in conjunction with the GES DAAC, and the system will be operated at the GES DAAC. NOAA will have the responsibility for the long-term archiving of the data and for the future NPOESS processing and archiving. Bob Murphy, NPP Project Scientist, talked about the NPP mission to the UWG. Additional information concerning NPP and NPOESS can be found in references listed at the end of this article.
- ***AVHRR NDVI land cover index production ends*** — This index production at the GES DAAC will stop after the data day of September 30, 2001, because the NOAA–14 polar orbiting satellite data quality is degrading. At launch its daylight local overpass time was 1:30 p.m.; on July 2, 2001, it was 5:04 p.m., and the westward drifting continues. The NDVI algorithm rejects as poor all observations near the sunrise and sunset terminators; thus, the NDVI is no longer available for many high latitude locations. The sensor calibration is also drifting. Some effort would be required to continue the NDVI time-series with data from the NOAA–16 satellite. The AVHRR on NOAA–16 has a split window visible channel and would require a new NDVI production algorithm.
- ***Charging for data*** — Historically the GES DAAC has distributed its data free of charge. However, with constrained resources and increasing data volumes and costs, NASA is planning to start charging for data some time in the future. For NASA derived data the fees will be set to cover the cost of filling orders. For data obtained from other sources the fees may vary depending on the specific arrangements made with the data sources. Specific details, including exact costs, bookkeeping procedures, and the date when charges will be imposed are still being worked out. For the NASA policy statement see

http://www.earth.nasa.gov/visions/data-policy.html

The UWG agreed that this was a reasonable policy. Several universities have such a policy for the data they supply to researchers and applied data users. Charging gives the archives additional funds to maintain and improve their services to customers and also discourages frivolous data requests. The UWG recommended that the GES DAAC announce several months in advance when charging would start and provide a clear schedule of fees. Researchers need to factor such costs into their research proposals, which often have to be submitted many months before the actual work is scheduled to begin.

A list of the members of the GES DAAC User Working Group can be found at

http://daac.gsfc.nasa.gov/DAAC_DOCS/UWG/

References

Earlier issues of the *Global Scanner,* newsletter of the GES DISC, are available on line at

http://daac.gsfc.nasa.gov/DAAC_DOCS/Newsletter/

Scroll down to Previous Issues and click the desired issue. In particular, see the related article "The GES DAAC, How It Works," in the the *Global Scanner,* Vol. 1, No. 1.

AVHRR Land Cover, NDVI Problems

http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/LAND_BIO/AVHRR_News.html

The GES DAAC User Working Group Web Site

http://daac.gsfc.nasa.gov/DAAC_DOCS/UWG/

See also

Kempler, S. J., and S. Morain. 2000. Goddard Earth Science (GES) Distributed Active Archive Center (DAAC) User Working Group (UWG); the minutes of the Nov. 2000 UWG meeting. In *The Earth Observer,* 13:1:19–24.

Earth Science Enterprise Web Page Data Policy

http://www.earth.nasa.gov/visions/data-policy.html

NPOESS Satellite Program

http://www.ipo.noaa.gov/

Ensuring the Climate Record From the NPP and NPOESS Meteorological Satellites

http://www.nas.edu/ssb/cdch2.htm

Processing the MODIS-Terra Consistent Year

continued from page 4

customers. Here "X" represents the rate at which data normally flow into the archive during nonreprocessing periods. Most customers will be interested in the higher level science products produced by MODAPS and in some regional Level 1B products. This constraint, while serious, still allows a wide distribution of data. The GES DAAC is working to maximize the number of customers served.

To help ensure a longer instrument lifetime MODIS was provided with a primary and a backup power supply and electronic systems (A & B). It ran on the A side electronics until the end of October 2000, then on October 30 it was switched to the B side electronics, which were preferred by the MODIS Science Team. On June 15, 2001, the B side power supply failed, and there may be other B side problems as well. The MODIS Engineering Team studied the problem carefully and on the evening of July 2 successfully put MODIS back into operation using the A side electronics and power supply. The MODIS Calibration Team is presently checking the data in order to optimize the A side calibration for the consistent year processing. Slight differences in the A and B side electronics cause their calibration tables to differ. While this problem won't affect the reprocessing program, it does cause some delay in the final processing of the MODIS Version 3 products for data taken after July 2, 2001. For recent information concerning the calibration lookup tables being used go to

http://mcstweb.gsfc.nasa.gov/product.html

The delivery, on July 18, of a new version of the ECS processing system that has overall control of the GES DAAC V2 processing and archive system did affect reprocessing for well over a month. Some problems arose in getting the new version to work as intended in the actual DAAC environment. Some incorporated, commercial off-the-shelf programs were upgraded about the same time. This combination of changes complicated the effort to get the changed ECS system to run as intended. One problem affecting reprocessing was that after the change it took much longer to pull data out of the tape archive and stage it for reprocessing. During this problem period the average reprocessing speed slowed from 1.4X to about 1.0X, while the forward processing by the ECS system was slowed even more.

In the near future, large data flows from Aqua and other future satellites will further increase the GES DAAC's load, thus additional expansions are being planned.

Acknowledgements

The authors thank Gary Alcott, Chris Lynnes, and Ed Masuoka for their input.

References

Earlier issues of the *Global Scanner,* newsletter of the GES DISC, are available on line at

http://daac.gsfc.nasa.gov/DAAC_DOCS/Newsletter/

Scroll down to Previous Issues and click the desired issue.

To check on or order MODIS data go to the GES DISC Home Page, Click on MODIS

http://daac.gsfc.nasa.gov/

Internet2 (I2) News, Abilene, and other information and links

http://www.i2x.org/

MODIS Home Page

A list of the MODIS Science Team members and many other items can be found here.

http://modis.gsfc.nasa.gov/MODIS/

MODIS PFM Instrument News

http://modis.gsfc.nasa.gov/MODIS/NEWS/news_PFM.html

MODIS Level 1B (Calibrated Radiances) Code and Product Information

http://mcstweb.gsfc.nasa.gov/product.html

New Data Products
General News
People in the News

as reported by
Lee Kyle, James Acker, Gregory Leptoukh, Jim McManus, Jianchun Qin, Hualan Rui, and Peter Smith

NEW DATA PRODUCTS & SERVICES

Detailed information about the archived data holdings at the GES DISC can be found at

http://daac.gsfc.nasa.gov

In this section we just emphasize important new happenings concerning our data holdings. These are arranged by data categories.

ATMOSPHERIC DYNAMICS

3-D dynamic and thermodynamic state of the Earth-atmosphere system, from satellite measurements and assimilation systems.

Item 1

Near real-time quality controlled NCEP observational data including global rawinsonde, aircraft, and surface synoptic reports are now available for display in our Online Analysis System (OASIS). Also, NCEP Aviation Forecast Model data will be accessible via OASIS for analysis time as well as 3-, 6-, and 9-hour forecasts. View these at

http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/atmospheric_dynamics/online_analysis /OASIS/html/

Item 2

A GrADS enabled, online gridded data image display tool is now available for atmospheric dynamics data sets. This allows users to visualize, manipulate, and assure data quality prior to acquiring data products. To use this online tool, go to

http://daacdev2.gsfc.nasa.gov/CAMPAIGN_DOCS/atmospheric_dynamics/online_analysis/CGI_GrADs/

HYDROLOGY

Global precipitation, its variability, and associated latent heating, important for studying the global hydrological cycle, climate modeling, and applications.

Item 1

A new TRMM WHOM data search and order mechanism, based on TRMM orbit numbers, was released in April 2001.

http://lake.nascom.nasa.gov/data/dataset/TRMM/01_Data_Products/01_Orbital/index.html

It provides users with a new interface and allows them to search and order multiple products by specifying orbit numbers.

Item 2

A new data set, GSSRB (Goddard Satellite-Retrieved Surface Radiation Budget in the Western Tropical Pacific), is available via the Goddard DAAC anonymous ftp at

ftp://lake.nascom.nasa.gov/data/TRMM/Ancillary/gssrb/

This sea surface radiation data set is produced by Dr. Ming-Dah Chou of NASA Goddard Space Flight Center (Code 913). It contains the downward solar (or shortwave) flux, downward thermal infrared (or longwave) flux, and upward longwave flux. The data set covers the domain 40°S-40°N and 90°E-170°W in the western Pacific and the period from January 1998 to April 2000 (no data from September to December 1999). The spatial resolution is 0.5° x 0.5° latitude-longitude, and the temporal resolution is 1 day.

Item 3

New versions of two data products available from the Goddard DAAC anonymous ftp site have recently been released.

3.1. GPCP Version 2 Combined Precipitation Data Set

ftp://daac.gsfc.nasa.gov/data/hydrology/precip/gpcp/gpcp_v2_combined/

GPCP (Global Precipitation Climatology Project) Version 2, comprising two products, the combined satellite-gauge precipitation estimate and the combined satellite-gauge precipitation error estimate, and the input and intermediate products, supersedes Version 1c.

3.2. GPROF 6.0 Gridded Orbit-By-Orbit Precipitation Data Set

ftp://lake.nascom.nasa.gov/data/TRMM/Ancillary/ssmi/

Version 6 of the GPROF (Goddard Profiling Algorithm) data set, providing instantaneous, gridded values of precipitation totals, based on SSM/I data, supersedes Version 4. This work was carried out as part of TRMM.

LAND BIOSPHERE

Long time-series vegetation and thermal infrared brightness temperature data sets for global change research.

Item 1

The AVHRR Pathfinder instrument's data quality is now highly questionable; consequently, the Goddard DAAC will cease production of this data set after September 30, 2001. More detailed information is available from

http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/LAND_BIO/AVHRR_News.html

Item 2

The original Pathfinder AVHRR Land (PAL) 8-km daily data have been reconfigured into tiled time-series files now orderable from the GES DISC DAAC at

http://daac.gsfc.nasa.gov/data/dataset/AVHRR/01_Data_Products/05_Tile_Products/

The original PAL interrupted Goode projection was divided into 40 x 18 "tiles" of 125 x 125 cells each. Cell sizes are 8 km, so each tile represents a 1000 x 1000 km region. There were originally 377 tiles containing at least one land cell, of which 322 tiles have been made available for ordering. The data are stored, by parameter, in monthly time-series files.

We hope this configuration of the PAL daily data will increase their accessibility. As an example of its utility, the original PAL daily data were archived in daily glo-

bal HDF files containing all 12 parameters. One of these files has an uncompressed volume of 228 MB. To process one parameter, such as NDVI, for a specific region over a 10 year period would require processing over 833 GB of data. The volume of NDVI data for a single tile, covering the same time period, would be 56 MB.

The original PAL daily data (1981 to 1994) have also been reprocessed, in the tiled time-series file configuration, correcting errors that existed in the original data set. Reprocessing involved calculating correct solar zenith angles and relative azimuth angles, and reextracting radiances from the original reflectance values of channels 1 and 2. The atmospheric corrections were then reapplied to the radiance producing new channel 1 and 2 reflectances. The NDVI was recalculated using the corrected reflectances.

A similar error existed in the NOAA-14 PAL data (1995 to 1999), except in this case the solar zenith angles were correct but were applied incorrectly to the normalization of channels 1 and 2. A variation of the method described above was used to correct this problem.

We hope the correction of the data and their reconfiguration into a more manageable form will open up the data to users who in the past did not have the resources to deal with the original daily data.

MODIS DATA SUPPORT

Radiance data and auxiliary information such as geolocation and cloud mask, atmospheric profiles, and higher level ocean color data.

Item 1

The MODIS Science Team has set up "Version 3" algorithms to produce improved MODIS science products (see this issue's feature article, "Processing the MODIS-Terra Consistent Data Products Year"). Version 3 MODIS products began public release on May 30, 2001. Some reprocessed data previous to November 2000 are available. The data products and data days available in Version 3 are continually updated as new products arrive. The MODIS Science Team is prioritizing the data periods being reprocessed to Version 3; thus, the reprocessing is not being done in strict chronological order. The latest information is found on the search and order site at

http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/MODIS/

Scroll down and click on "Data Access"

Note: The Version 3 products for the period July 2, 2001, and following are not at present optimal (see Item 2 below).

Item 2

There is a permanent MODIS data gap from June 15 to July 2, 2001, because of an instrument problem. The MODIS instrument is provided with a primary and a backup power supply and electronic systems (A & B). It ran on the A system until the end of October 2000 when it was switched to the B system, which was preferred by the MODIS science team. On June 15, 2001, the B power supply failed, and there may be other B system problems as well. After a careful study of the problem the MODIS engineering team put MODIS back into operation using the A system. This switch required revised calibration lookup tables (LUT) for data taken after the July 2 turn on. New provisional A system LUTs were delivered by the calibration team on July 24. Improved A system LUTs are to be delivered soon. For additional and update information see

http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/MODIS/rad_geo/MOD01_L1A_disclaimer.s html

OCEAN COLOR

Remote sensing ocean color data used to investigate ocean productivity, marine optical properties, and the interaction of winds and currents with ocean biology.

Item 1

Our latest Science Focus page is entitled "SeaWiFS and Global Warming." Another recent Science Focus page of interest is "Ethiopia, the Red Sea, and the Nile River." The next two currently in production will be about the Bering Sea and the influence of turbidity on ocean color data. The Bering Sea page should be made public the first week of September. For these and other interesting subjects see

http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/OCDST/science_focus.html

Item 2

The SeaWiFS Data External Browser will release Version 1.0 in early autumn. This version adds a basic search engine and a simpler ordering interface than the beta test version. Browse images for all files archived from the beginning of the mission (September 1997) through June 2001 will be available on a set of CDs. When the browser is near completion, an announcement will be sent to the ocean color mailing list to determine initial demand.

Researchers may also find useful products in our other important data set collections.

ATMOSPHERIC CHEMISTRY

Ozone and other trace gas compositions, dynamics, and energy interactions of the upper atmosphere.

FIELD EXPERIMENTS

Aircraft and ground based measurements of meteorological variables designed to improve science algorithms and validate satellite-derived data products.

INTERDISCIPLINARY

Global land, ocean, and atmospheric parameters mapped to uniform spatial and temporal scales for basic research and applications studies.

For more details about the GES DISC data holdings and to order data see our Home Page or contact us by eMail, phone, or fax.

http://daac.gsfc.nasa.gov/

For MODIS User Services
eMail: daac_usg@gsfcsrvr4.gsfcmo.ecs.nasa.gov
phone: 301–614–5473
fax: 301–614–5304

For other products' User Services
eMail: daacuso@daac.gsfc.nasa.gov
phone: 301–614–5224 or 1–877–422–1222
fax: 301–614–5304

GENERAL NEWS

DISC Happenings

Aqua

Preparations continue for the GES DAAC's support of the data to come from the MODIS and AIRS-HSB-AMSU instruments on the Aqua satellite, which is scheduled to be launched no earlier than December 2001.

Future Program

The GES DISC was selected to be the site for the NPP science data segment Level 1 processing and calibrations systems. Planned for launch in 2005, the NPP satellite will continue the climate data records

continued on page 10

coming from the Terra and Aqua satellites. It will also prototype instruments and processing systems for future NOAA polar orbiting environmental satellites. It is a joint NOAA-NASA project (see discussion in this issue in "The GES DAAC User Working Group" article).

Awards

In May three awards were made to the GES DISC staff that collectively covered all personnel. They were presented by Al Diaz, Director of the Goddard Space Flight Center. The awards are as follows:

Quarterly Customer Service Excellence Award — awarded to the GES DISC Data Support and Operations Groups led by George Serafino and Gary Alcott.

> In recognition of your contributions in achieving an extremely high level of data support and operations reliability by exceeding user expectations.

Quarterly Outstanding Teamwork Award — awarded to the Engineering and Science Integration Teams led by Chris Lynnes and Bruce Vollmer.

> In recognition of your outstanding contributions to integrate new system functionality, science software, and advanced technologies into existing data management production systems that greatly facilitate the achievement of NASA's scientific goals.

Excellence in Information Science and Technology Award — GSFC's highest information science award, to Chris Lynnes (see his personality sketch in the People section).

Future Summer Student Mentoring

Steve Wharton, Chief of the Global Change Data Center, has asked Peter Smith to act as the focal point and liaison for future summer student mentoring opportunities within the data center, which includes the GES DISC. This summer Matt Mitchell, the student mentored by Peter and his AVHRR Data Set Team, won an award for his work here (see People). Peter strongly encourages other data set leads to step up to the plate and volunteer to be a mentor.

Publications and Presentations

AGU Spring Meeting — The GES DAAC had an information booth and made several presentations at the American Geophysical Union (AGU) meeting in Boston, May 29–June 2, 2001. Some of the poster presentations included hands-on demonstrations. Abstracts of these presentations can be found in *2001 Spring Meeting, May 29–June 2, 2001,* Section U, Session U21A; Published as a supplement to *EOS, Transactions, AGU,* Vol. 82, No. 20, May 15, 2001, and include

Web Based Hierarchical Ordering Mechanism (WHOM) Tool for MODIS Data From Terra, Mohammed S. Sikder, Peggy Eaton, Gregory Leptoukh, Nancy McCrimmon, & Bryan Zhou.

TERRA/MODIS Data Products & Data Management at the GES-DAAC, Awdhesh K. Sharma, S. Ahmad, P. Eaton, J. Koziana, G. Leptoukh, D. Ouzounov, A. Savtchenko, G. Serafino, M. Sikder, & B. Zhou.

Simple Mapping Tools From the Goddard DAAC Earth Sciences MODIS Data Support, Andrey K. Savtchenko.

Web Based Tools for Checking EOS Data Integrity, Krishna Tewari, A. K. Sharma, L. Fenichel, S. Kreisler, G. Leptoukh, C. Lynnes, G. Roth, & R. Strub.

Online Analysis of BUFR-Formatted Quality-Controlled NCEP Final Observation Data at the Goddard Earth Sciences DISC/DAAC, Carrie S. Phelps, Jianchun Qin, Mahabaleshwara S. Hegde, & Donald R. Frank.

Session B51A: *MODIS Ocean Color, SST, & Primary Productivity Products for Regional & Global Studies,* James Koziana, Wayne Esaias, Gregory Leptoukh, Andrey Savtchenko, and George Serafino.

Workshop for Earth Science Satellite Remote Sensing Data Processing, Analysis and Applications, at George Mason University, Fairfax, VA, June 4–12, 2001 — GES DISC personnel presented several lectures.

Jim Acker, *SeaWiFS Products & Applications.*

Long Chiu, *TRMM Products & Applications.*

Greg Leptoukh, two lectures, *MODIS Data Processing* and *MODIS Data Searching & Ordering From GDAAC.*

Chris Lynnes, two lectures, *MODIS Direct Broadcast System* and *MODIS Simple & Scalable Script-Based Science Processor (S4P).*

John Qu, two lectures, *TOMS Products & Applications* and *MODIS Data Subsetting & Image Processing.*

EDG Workshop at GSFC, June 13–15 — Peggy Eaton presented the GDAAC Report.

IGARRS — MODIS Data Support members presented several talks and posters at IGARSS'01 in Sydney, Australia, July 9–14, 2001.

Jim Koziana, poster: *Ocean Data From MODIS at the NASA Goddard Earth Sciences DAAC* and oral: *MODIS Atmospheric Data Products at the Goddard Earth Sciences DAAC.*

Greg Leptoukh, poster: *MODIS Data Ingest, Processing, Archiving, & Distribution at the Goddard Earth Sciences DAAC.*

John Qu, poster: *New Application of Earth Science Remote Sensing Data at NASA/GES DISC: Remote Sensing Information Partner (RSIP) With Rutgers University.*

Steve Berrick, oral: *Simple, Scalable, Script-Based Science Processor (S4P).*

William Teng, Nathan Pollack, George Serafino, Long Chiu, and Paul Sweatman, poster: *GIS & Data Interoperability at the NASA Goddard DAAC.*

Long Chiu, George Serafino, and William Teng, poster: *Applications of TRMM Data.*

PEOPLE IN THE NEWS

Personality Sketch: Christopher (Chris) Lynnes, leader of the GES DISC System Engineering Group

Chris Lynnes earned a Ph.D. in geophysics from the University of Michigan in 1988. After 3 years in nuclear test detection research, he joined the Goddard Distributed Active Archive Center (DAAC), first as contractor software lead, then as civil servant and systems engineer in 1994. This spring he received the Excellence in Information Science and Technology (IS&T) Award, GSFC's highest information science award. GSFC Director, Al Diaz, made the presentation in a ceremony on May 23, 2001. The award was in recognition of Chris's contributions to the GES DAAC; its citation summarizes some of his major accomplishments.

> Dr. Lynnes's major IS&T contributions are the systems whose design and construction he led on behalf of NASA Earth science research applications, including the Distributed Active Archive Center's (DAAC) first operational system, the Tropical Rainfall Measuring Mission (TRMM) Support System, the Web Hierarchical Ordering Mechanism (WHOM), and the Simple Scalable Script-Based Science Processor (S4P), which was a major contributor to the successful and timely delivery of MODIS data products. These contributions, plus his expertise in applying state-of-the-art IS&T to Earth science needs, have been truly exemplary and internationally acknowledged.

Chris has been a dynamic leader of the DAAC system engineering team. As mentioned in the award citation, he led them in the creation and maintenance of GES DAAC's V0 (UARS, SeaWiFS, etc.) and V1 (TRMM) data archiving systems. He redesigned the GES DAAC WWW archi-

tecture for searching and ordering data to make it easier to use and simpler to maintain and modify. WHOM is a principle example of this work. Starting in 1995 he led the development of a Web site (http://EarthInteractions.org) to host Earth Interactions, a peer reviewed electronic journal published by the American Meteorological Society, American Geophysical Union, and Association of American Geographers. The journal went on line in 1997 and has four yearly volumes on the Web.

The S4P processor, whose development he guided, is a vital element of the GES DAAC's system. Chris has consulted on the development of the EOSDIS Core System (ECS) since 1995, and he continues to develop extensions to enhance the operation of ECS at the Goddard DAAC. Originally S4P was designed as a fall back system in case the contracted ECS system wasn't ready by the time the Terra satellite launched. It wasn't needed for this, but it came in handy when it was decided to develop a Goddard system to process direct broadcast MODIS data. Local users can receive these data as the satellite passes overhead, but a processing system is needed to produce usable products. The Goddard direct broadcast system became operational in spring 2000. Since then, variations of the S4P system have been set up to handle quick response MODIS processing for those needing the observations quickly, and presently a S4PM (MODIS) is handling the reprocessing of Terra MODIS data.

To illustrate the simplicity and flexibility of the S4P design, Chris tells the story of a quick response request to support teams fighting forest fires in and around Montana in late August 2000. Members of the MODIS science team agreed to furnish prompt MODIS derived fire location information for that region to the Forest Service. The fire locations were in turn dependent on Level 1B data from the GES DAAC. As the DAAC MODIS L1 stream is normally several days behind real time and the arrival of the L0 data, Chris originally planned to run the MODIS algorithms manually on individual L0 input files as soon as they were acquired by the ECS V2 system. Over Labor Day weekend, he would simply put in subscriptions for the L0 input data and log in remotely to run the algorithms when the data arrived. However, on arriving home that Friday, his wife informed him that the phones were out. Faced with the prospect of camping in his office all weekend to wait for incoming Level 0 data, he spent 6 hours implementing an S4P-based system to automatically detect the data arrival, process the correct time period within the file (covering Idaho and Montana), and push the data to the MODIS team where it was turned into fire locations.

Chris is not resting on his laurels but is continuing to work to make the GES DISC a technically superior and user and operator friendly system and organization. He has been a frequent contributor to this publication and more details concerning some of his efforts at the GES DAAC can be found in articles in previous issues of this newsletter.

Developing a simplified MODIS processing system, *Global Scanner,* Vol. 2, No. 3-4.

Data mining at the GES DAAC, *Global Scanner,* Vol. 2, No. 2.

MODIS L1 data & L1 direct broadcast now available, *Global Scanner,* Vol. 1, No. 4.

WHOM, a Web hierarchical ordering mechanism, *Global Scanner,* Vol. 1, No. 1.

http://daac.gsfc.nasa.gov/DAAC_DOCS/Newsletter/

Scroll down to Previous Issues and click on the desired issue.

Peter Smith, lead of the AVHRR Data Support Team, reports that Matt Mitchell won a Summer Student Award

This year the AVHRR data set team participated in the summer student ACCESS program managed by the Equal Employment Opportunity Office (EEO), Code 120. Our student was **Matthew Mitchell** from the University of Northern Illinois, Michigan. Matt is wheelchair bound because of a motorcycle accident that occurred when he was 22. He was hit head on by a pickup truck and sustained extensive spinal injuries causing him to lose the use of his legs and severely limiting the movements of his arms. It took him about 10 years to come to terms with this state of

continued on page 12

cut here

If you wish to receive *The Global Scanner* each season, fill out this form and mail or fax it to us.

CHECK ONE: ____ I am a new subscriber. ____ I am already on the list; this is an update.

NAME _____

ORGANIZATION _____

ADDRESS _____

_____

_____

CITY _____ STATE _____ ZIP _____

PHONE _____ FAX _____

E-MAIL _____

affairs. He had always assumed that he would take over the running of his father's dairy farm and now this was out of the question. He finally got himself straightened out and made the decision to go back to school and obtain a science degree. He is now 39 years old and has completed his junior year and this fall will be a senior. He has maintained a GPA of 4.0 throughout his 3 years of college.

Matt spent 10 weeks with us and was put to work on developing a Web based analysis tool for the AVHRR subsetted continental data that we host on our FTP site. To develop this tool, Matt had to learn PERL, C, and IDL and bone up on his existing UNIX and HTML skills. Although familiar with HTML, he had to extend his Web programming knowledge to cover forms processing using the Common Gateway Interface (CGI). Having gotten all those ducks in a row he then built an analysis tool with which a Web user can select a given geographic spatial area and temporal period and create a time series plot of up to 5 AVHRR parameters (NDVI, Channels 1 and 2 reflectances, Channels 4 and 5 Black Body Temperatures). On-the-fly time-series GIF images of means and standard deviations for each parameter for the spatial area of interest are generated using IDL and displayed to the user on the final Web page. During the process of developing this tool Matt had to make several adjustments to improve its performance and user friendliness.

The final week of the program calls for each student to present his project to a panel of three judges and the student's division chief. Steve Wharton, our division chief, attended Matt's presentation and was extremely impressed with the content and smoothness of Matt's talk. So apparently were the judges, as Matt was selected with 9 other students from a group totalling about 70 students to receive the Rahsaan Jackson award.

I encouraged Matt to interact with all members of the AVHRR team and to learn as much about how we go about maintaining and producing a remotely sensed data product. He was very interested and eager to learn about all facets of our data set responsibilities. In fact, he got so fired up that he is strongly considering applying to enter the student Coop program at Goddard where a student splits his time equally between work and attending college for a graduate degree with college tuition being paid for by NASA. Unfortunately 902 has limited openings and Matt is being strongly encouraged to apply to code 586 where there are a larger number of employment opportunities.

In summary, the experience and benefits accruing from mentoring a visiting student are extremely positive. It is, of course, a two way street, and much depends on the quality and attitude of the student. I was lucky and got a student who didn't mind hard work and was eager to learn as much as possible. I think Matt got a lot out of this experience and the AVHRR team in return also profited greatly. Steve Wharton has asked me to act as the focal point and liaison person for future summer student mentoring opportunities within 902, and I will be strongly encouraging other data set leads to step up to the plate and volunteer to be a mentor.

Well done Matt, and congratulations on your award!

— Peter Smith

The Global Scanner is a publication of NASA's Goddard Space Flight Center Earth Sciences (GES) Data & Information Services Center (DISC).
Send comments or articles or information you wish to submit to the Editor: lkyle@daac.gsfc.nasa.gov. Editors: Lee Kyle, Diana Sunday, Steve Kempler.
Help Desk: daacuso@daac.gsfc.nasa.gov — Phone: 301-614-5224 or 1-877-422-1222.

the global scanner
Goddard Earth Sciences Data & Information Services Center
Code 902
NASA Goddard Space Flight Center
Greenbelt, MD 20771